

**PATENT ABSTRACTS OF JAPAN**

(11)Publication number : 2003-006642

(43)Date of publication of application : 10.01.2003

---

(51)Int.Cl.

G06T 7/00

G06T 1/00

G06T 7/20

H04N 7/18

---

(21)Application number : 2001-189368 (71)Applicant : SHIMIZU TOSHIHIKO

MINAMI MAMORU

(22)Date of filing : 22.06.2001

(72)Inventor : MINAMI MAMORU

SHIMIZU TOSHIHIKO

---

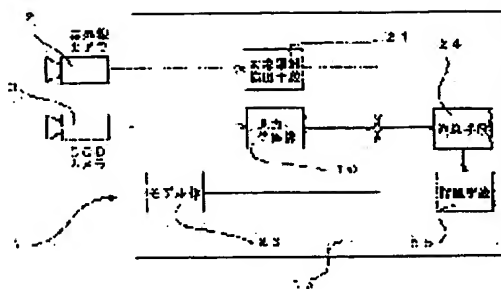
(54) IMAGE PROCESSOR UTILIZING THERMAL IMAGE

from 21.12.01

## (57)Abstract:

PROBLEM TO BE SOLVED: To narrow down an inspection object area at high speed while utilizing a thermal image and to exactly detect a human being without being affected by use conditions or his/her looks.

SOLUTION: This device is provided with an object area detecting means 21 which has an infrared camera 2 for sampling a thermal image 12 and a CCD camera 3 for sampling a visible image 10 within almost the same area for finding an inspection object area 20 of a heating body position from the thermal image captured by the infrared camera 2, a model body 23 composed of a multiplex structure



area 22 for making the image object into simplified model, an arithmetic means 24 with which a plurality of model bodies 23 are overlapped within the inspection object area 23 of an input raw image 10 by the CCD camera 3 for applying pixel arithmetic within the inspection object area 20 while using the raw image 10, and evaluating means 25 for evaluating the arithmetic result.

---

## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against

examiner's decision of rejection]

[Date of extinction of right]

**\* NOTICES \***

**JPO and NCIP are not responsible for any damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. \*\*\*\* shows the word which can not be translated.

3. In the drawings, any words are not translated.

---

## CLAIMS

---

[Claim(s)]

[Claim 1] Are the image processing system which performs human detection using thermal imagery, and it has an infrared image input means to sample thermal imagery, and an image pick-up means to picturize the visible image in an abbreviation same field. An object domain detection means to ask for the inspection object domain of a hot object location from the thermal imagery caught with the infrared image input means, The model object which consisted of multiplet-structure fields and carried out simple modeling of the image object, The image processing system using the thermal imagery characterized by for this model object piling up to the above-mentioned inspection object domain in the input student image caught with the image pick-up means, providing an operation means to perform a pixel operation using the pixel brightness value in the field, and an evaluation means to evaluate the above-mentioned result of an operation, and detecting human being.

[Claim 2] The image processing system using the thermal imagery according to claim 1 with which magnitude differs, and a model object stands in a row in the abbreviation vertical direction, and consists of double structure fields, and changes.

[Claim 3] The image processing system using the thermal imagery according to claim 2 which the lower part field of the model object which stands in a row up and down is established in dimension size, and 2 \*\*\*\*s of the double structure field are further carried out in the vertical direction, and changes.

[Claim 4] The image processing system using the thermal imagery according to claim 2 with which the model object which magnitude differs and stands in a row in the abbreviation vertical direction stands in a row in the shape of three step.

[Claim 5] The image processing system using thermal imagery claim 1 which calculates the absolute value of the difference of the brightness value sum total of a pixel with which an operation means corresponds to the outside field of a model object, and the brightness value sum total of the pixel applicable to the inside field of a model object, 2, 3, or given in four.

[Claim 6] The image processing system using thermal imagery claim 1 to which an object domain detection means catches change of the location of the hot object acquired with the infrared image input means, and detects an inspection object domain, 2, 3, 4, or given in five.

[Claim 7] The image processing system using thermal imagery claim 1 which responds to the magnitude of the inspection object domain of the hot object acquired with the infrared image input means, expands or contracts, and changes the dimension of a model object, 2, 3, 4, 5, or given in six.

[Claim 8] The image processing system using thermal imagery claim 1 which makes the image field obtained by carrying out a zoom in the direction of the maximum upper part field section of a model object when it detects that an evaluation result is human being a new inspection object domain, and performs \*\*\*\* detection of human being with a model object with a still newer multiplet-structure field, 2, 3, 4, 5, 6, or given in seven.

[Claim 9] The image processing system which used thermal imagery claim 1 which records an image pick-up image, 2, 3, 4, 5, 6, 7, or given in eight when it detected that an evaluation result is human being or its face section.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the image processing system which detects human being using thermal imagery.

[0002]

[Description of the Prior Art] Although the thing to which an image processing system generally makes binary the input student image (input image which has brightness gradation) sampled with the CCD camera from the former, perform thinning etc., extract the geometric focus, such as an endpoint, the branch point, or a configuration, and objective image recognition is made to perform is known In this case, binary-ized

processing is the comparison with threshold level, and it is not the configuration object geometrically specified from making missing the information which the image itself has originally [ — data are rounded off binary / of 0 or 1 / — ]. Detection of the focus was difficult for recognition of human being who changes the appearance free and moves, and unsuitable for it.

[0003] Then, a changed part of the monitor image by the CCD camera arranged on admission port etc. is caught in an image processing, the focus is extracted and carried out [ human ], a kimono is detected, an infrared detector is further used together and the approaches (JP,2000-76521,A, JP,2000-155177,A, JP,2000-209573,A, etc.) of making it judge that the focus is human being are proposed.

[0004] On the other hand, approaches, such as JP,2001-92972,A which does not perform binary-ized processing but performs image recognition with GA (heredity algorithm)-heuristics, are learned in recent years.

[0005] The image recognition approach of this JP,2001-92972,A distributes over an input image the search model which uses as a model the luminance distribution and the simple configuration of the object which has become clear beforehand, is evolving the gene information which shows the location of a search model using GA, and outputs the positional information of an object.

[0006] When distribution of the brightness value of a search model and the distribution of the brightness value of an input image of the process of evolution of the gene by this GA corresponded, it evolves the gene (location) information on a search model to the location where a correlation value becomes high using the correlation function using the principle to which a correlation value becomes high (a coincidence location shows a peak), and it would acquire the solution.

[0007] In addition, in the above-mentioned Prior art, it was what processes whether the focus of a dynamic image is extracted from a binary-ized image, or a search model is evolved by GA.

[0008]

[Problem(s) to be Solved by the Invention] Since threshold level is influenced by the above-mentioned conventional binary-ized image in the condition of lighting etc. in the thing of a configuration of extracting the focus of a dynamic image, it is weak to disturbance and there is risk of losing the information which it should have essentially by the slight roundness of the information by binary-izing. Moreover, it is the lobby with much receipts and payments of people which inserts even when sunlight etc. is indirect (the quantity of light changes in day ranges and Nighttime), and the most difficult work extracts the focus of human being from whom appearances, such as

makeup and a hairstyle dress, moreover differ and who is moving.

[0009] Therefore, if it is in things, such as JP,2000-76521,A, JP,2000-155177,A, and JP,2000-209573,A, it is regarded as human being with the output of the animation information that some objects moved on parenchyma, and an infrared detector, and is hard to say that human being is not necessarily detected correctly.

[0010] And if it is in the approach of extracting such an animation From it being extracted as the focus to which only the animation section of only a bag gives priority, and being judged, when the bag which people have is shaken although human being existed in the image, in spite of having said that the detection could not be performed and having performed human detection -- having been undetectable (that is, there being no people) -- the problem which commits the serious mistake to say is pointed out. Then, in order to avoid it, it was the thing which makes human being's existence judge with the output of an infrared detector in addition to the object which moves.

[0011] Therefore, it is the place which is the area (admission port) of a certain fixed fixed background, and is not considered as use range except people's penetration. It is the available image-processing approach only in the situation which change (animation) of an image tends to catch. The thing of an application for which another side (for example, the CCD camera itself) samples the image of swing movable within the limits for a neck, If it does not detect an application which is installed in inside of a shop, a supermarket, etc. with much fluctuation of the lighting used as a noise and receipts and payments of people, or that he is human being correctly, there is a trouble that it cannot respond in the application to affect a human life.

[0012] On the other hand, if it was in the object of a proper in the thing of the above-mentioned conventional configuration which performs retrieval by GA, there was no rounding-off of the information by binary-izing, it was strong to disturbance, moreover, unlike the thing of the comparison by pattern coincidence, there was a merit which can look for the unknown object of the whereabouts, and it was an effective approach as compared with the conventional thing.

[0013] However, the thing of an approach which searches with this search model was what has a trouble for it to be necessary to prepare beforehand the search model with which distribution of a brightness value has become clear, and detect human being.

[0014] namely, the case which detects common unspecified human being although it can be searched for it if it becomes where a specific individual's appearance has become clear beforehand -- people -- the color of the skin, and the color of hair -- a race -- differing -- moreover -- makeup -- carrying out -- in addition -- and since it is clothed in various dresses, it is difficult absolutely to model distribution of the

brightness value.

[0015] And there is a fault depended on thinning and the focus extract by binary-izing that retrieval will take much time amount like a thing, and there is a problem in arresting much human being quickly.

[0016] In addition, in JP,2001-92972,A, if it is in this approach, needless to say, it is difficult [ it / in order to give real time nature the example which makes an optimum solution a thing with the highest correlation value to desired control timing is shown, without waiting for convergence of evolution, but ] to pinpoint human being's location correctly.

[0017] Even if this has a factor depending on a contingency in the process of the method of rose firewood of a search model, or evolution and the candidate of a solution becomes, a thing with the not necessarily highest correlation value to desired control timing is because a true solution does not become.

[0018] Furthermore, since many locations of brightness value distribution where a correlation value shows a middle peak also existed in an image even if it is a thing besides an object, GA might perform gaze actuation to the temporarily mistaken field, and it had become a problem certainly making high-speed processing perform moreover.

[0019] Then, as an image processing system which performs human detection, what is moreover detected certainly at high speed was desired, without being dependent on factors, such as an operating condition and people's appearance.

[0020] This invention aims at offering the image processing system which detects human being correctly, without being confused by an operating condition, people's appearance, etc. while it narrows down an inspection object domain at high speed using thermal imagery.

[0021]

[Means for Solving the Problem] In order that the image processing system using the thermal imagery of this invention may solve the trouble of the above-mentioned conventional example, Are the image processing system which performs human detection using thermal imagery, and it has an infrared image input means to sample thermal imagery, and an image pick-up means to picturize the visible image in an abbreviation same field. An object domain detection means to ask for the inspection object domain of a hot object location from the thermal imagery caught with the infrared image input means, The model object which consisted of multiplet-structure fields and carried out simple modeling of the image object, It is characterized by for this model object piling up to the above-mentioned inspection object domain in the

input student image caught with the image pick-up means, providing an operation means to perform a pixel operation using the pixel brightness value in the field, and an evaluation means to evaluate the above-mentioned result of an operation, and detecting human being.

[0022] Moreover, it is suitable that magnitude differs, and a model object stands in a row in the abbreviation vertical direction, and consists of double structure fields, and changes.

[0023] And if it is in any of whether the field of the model object which the lower part field of the model object which stands in a row up and down is established in dimension size, and whether 2 \*\*\*\*s of the double structure field are carried out in the vertical direction, and it grows into it further differs from magnitude, and stands in a row in the abbreviation vertical direction stands in a row in the shape of three step they are, exact detection is performed, without being dependent on people's appearance.

[0024] And calculating the absolute value of the difference of the brightness value sum total of a pixel with which an operation means corresponds to the outside field of a model object, and the brightness value sum total of the pixel applicable to the inside field of a model object produces a difference with a background by the easy operation.

[0025] Moreover, that to which an object domain detection means catches change of the location of the hot object acquired with the infrared image input means, and detects an inspection object domain is suitable.

[0026] Detection of a child, an adult or a distant place, or near can be performed by responding to the magnitude of the inspection object domain of the hot object furthermore acquired with the infrared image input means, and expanding, or reducing and changing the dimension of a model object.

[0027] In addition, if the image field obtained by carrying out a zoom in the direction of the maximum upper part field section of a model object is made into a new inspection object domain and it is adapted in a model object with a still newer multiplet-structure field when it detects that an evaluation result is human being, \*\*\*\* detection of human being can be performed. And if the image pick-up image is recorded, the image which specifies the person will be obtained.

[0028] (Operation) This invention can perform the following operations by the above-mentioned configuration. By namely, the thing searched for for the inspection object domain of a hot object location from the thermal imagery caught with the infrared image input means Thermal imagery Human being (part which the skins, such as a face and a hand, have exposed correctly), an animal, Certainly, it can grasp,



moreover, image retrieval can be caught by ultra high-speed, without reacting to a heating element, a heat-regenerative element, etc., and detecting the location appropriate for existence of human being, and some inspection object domains can be quickly set up from the distribution situation of a heating element in consideration of human being's magnitude.

[0029] The model object which consisted of multiplet-structure fields and, on the other hand, carried out simple modeling of the image object models the configuration of an object in simple, and is equipped with the field which serves as the multiplet structure (for example, double structure) further. And in order to make it judge from the brightness value of the pixel which laps with this multiplet-structure model that he is human being, the field inside double structure is made to correspond to the human being itself, and that by which an outside field is equivalent to a background is shown. In addition, since there is no need for retrieval using search models, such as GA as shown in the conventional example by detection of the above-mentioned heating element, on this model object, the distribution data of the brightness value of an object may be unnecessary, and you may be homogeneity data of an arbitration brightness value, and may be \*\*\*\* which defines only a configuration.

[0030] Then, \*\*\*\* of a model object piles up one by one into the inspection object domain called for the account of a top to the input student image by the image pick-up means, and a pixel operation is performed in an inspection object domain from the pixel which has a brightness value. The heavy activity of this model object is repeated one by one by the arrangement which can insert a model object into the limited narrow inspection zone, and it is with the field of the above-mentioned inside (human being), and an outside (background) field. If the brightness value of the pixel in each field is calculated and a difference arises in the brightness value of an outside field (background), and the brightness value of an inside field (human being), it can know existence of the body which agrees in the shape of a model bodily shape with this difference, as people identify an object.

[0031] an inside field constitutes human being's configuration, and since it is moreover a hot object, this result of an operation is the difference of the result of an operation in the case — if it is \*\*, it can be estimated that it is human being.

[0032] Moreover, if magnitude differs, and a model object stands in a row in the abbreviation vertical direction, and consists of double structure fields and changes, what the upper field of a model object shows can be used as a head, and what a downward field shows can be made into a soma. In these two fields, respectively, by the relation between an inside field (head), an outside field (background) and an inside

field (soma), and an outside field (background), if a difference is produced in a brightness value in each field, some which emit heat are there, and there is a configuration object which is moreover equivalent to the head and the body, namely, it can be judged that he is human being.

[0033] And if the lower part field of the model object which stands in a row up and down is established in dimension size and 2 \*\*\*\*s of the double structure field are further carried out in the vertical direction, what the divided up field shows can be made into a drum section, and what a lower field shows can be used as the leg. Under besides, by the relation between an inside field (drum section), an outside field (background) and an inside field (leg), and an outside field (background), it can judge that a drum and a foot are there if a difference is produced in a brightness value, and the above in two fields serves as the same a model object which performs exact detection, without being dependent on people's appearance, if it is such a model object.

[0034] If the field of the model object which magnitude differs and stands in a row in the abbreviation vertical direction on the other hand is one of those which stand in a row in the shape of three step, what the up field divided into three steps shows can be used as a head, what the CHUBU ENGINEERING CORPORATION field shows can be made into a drum section, and what a lower field shows can be used as the leg. It can judge that a drum and a foot are there if a difference is produced in a brightness value like the thing of the above-mentioned 2 division field in this Nakashita 2 \*\* field in the relation between an inside field (drum section), an outside field (background) and an inside field (leg), and an outside field (background), and it becomes the model object which performs exact detection, without being dependent on people appearance like the above, if it is such a model object.

[0035] in addition -- if it is what calculates the absolute value of the difference of the sum total of the brightness value of each pixel with which an operation means corresponds to the outside field of a model object, and the sum total of the brightness value of each pixel applicable to the inside field of a model object as mentioned above -- this double structure -- therefore, it is in the location where a model object laps with an object configuration only by the operation of easy addition and subtraction in the case -- a \*\* difference (background) can produce.

[0036] Moreover, though the mannequin etc. was heated and placed, for example when an object domain detection means catches change of the location of the hot object acquired with the infrared image input means and detected the inspection object domain, in order not to move, it does not become a candidate for detection, but by the

heating element which does not move similarly being disregarded, there is no futility and only human being can be detected more certainly.

[0037] It responds to the magnitude of the inspection object domain of the hot object furthermore acquired with the infrared image input means, and if the dimension of a model object is expanded, or reduced and changed, human detection can be made to perform in more detail using the model of a child, an adult or a distant place, or near.

[0038] In addition, if the image field obtained by carrying out a zoom in the direction of the maximum upper part field section of a model object (head) is made into a new inspection object domain and it is adapted in a model object with a still newer multiplet-structure field when it detects that an evaluation result is human being that by which the new multiplex model equivalent to people's head, the face section, etc. is adapted only for the limited maximum upper part field section circumference, and it is detected that he is already human being -- in addition -- and in the limited image field by which a zoom was carried out (with for example, model with multiplex fields, such as a difference of a face and hair) It can evaluate by performing the pixel operation by the brightness value like the above using the model object of only the face section. Therefore, when the target face is suitable in the direction of an image pick-up means only in the limited zoom field, without carrying out the retrieval search of all the input images, it becomes possible to catch the zoom image of the face section.

[0039] Furthermore, to the timing which catches the above-mentioned \*\*\*\*, if the image pick-up image near the face section model object location is recorded serially, it becomes what obtains the image which specifies the person, and the image contributed not only to human detection but to person specification can be recorded.

[0040] Moreover, the image processing system using the thermal imagery of such this invention can detect human being certainly at high speed, without being dependent on factors, such as an operating condition and people's appearance.

[0041] Moreover, the image processing system which catches the face zoom image for specifying detected human being is realizable.

[0042]

[Embodiment of the Invention] The gestalt of operation of this invention is briefly explained with reference to drawing 1 - drawing 17 .

[0043] The image processing system 1 using the thermal imagery of this invention It is image processing system 1a which detects human being 30 using the image (thermal imagery) caught with the infrared camera 2. It has CCD camera (image pick-up means) 3 of optical black and white which can picturize the part or the same visible image (image of the light) 10 in the infrared camera (infrared image input means) 2 which

samples thermal imagery 12, and an abbreviation same field. An object domain detection means 21 to make binary thermal imagery 12 caught with the infrared camera (infrared image input means) 2, and to decide the magnitude of the inspection object domain 20 from the coordinate location of the hot object object 13 more than fixed \*\*, The model object 23 which consisted of double structure-like ellipse fields 22, and modeled the configuration of an image object beforehand in simple, The above-mentioned model object 23 piles up regularly (or irregularly) into this inspection object domain 20 of the input student image (input image which has brightness gradation) 10 by CCD camera (image pick-up means) 3. An operation means 24 to perform a pixel operation from the brightness value of the pixel in the field which the model object 23 shows, and an evaluation means 25 to evaluate by whether the above-mentioned result of an operation reaches the value set up beforehand are provided, and human being 30 is detected.

[0044] moreover, the dimension which is equivalent to an intersection and a head 31 in that the model object 23 hits a man's 30 neck (K points) — smallness — the dimension equivalent to the 1st field and soma 32 — size — it consists of structures of the elliptical-ring-structure field 22 of a duplex so that magnitude may differ, and it may stand in a row mostly in the abbreviation vertical direction and it may gather at a top Norikazu point (K points) with consisting of the 2nd field.

[0045] And the lower part field of the model object which stands in a row up and down is established in dimension size, 2 \*\*\*\*s of the double structure field (the 2nd field 32) are made still more equal by abbreviation with the vertical direction, and the soma 32 (the 2nd field) is divided into drum section 32A (field of the 2nd A), and leg 32B (field of the 2nd B). In addition, even if it carries out a basis, it becomes the thing which stands in a row the model object field which magnitude differs and stands in a row in the abbreviation vertical direction in the shape of three step and which can be divided into a soma 32 (the 2nd field) and a foot 33 (the 3rd field) similarly.

[0046] The operation means 24 is multiplying by it and asking for the weight value the absolute value of the difference (  $|J \cdot \sigma_{Sj} - I \cdot \sigma_{Si}|$  ) of the brightness value sum total (  $J \cdot \sigma_{Sj}$  ) of the pixel applicable to the outside field  $S_j$  of the double structure field 22 which constitutes the model object 23, and the brightness value sum total (  $I \cdot \sigma_{Si}$  ) of the pixel applicable to the inside field  $S_i$  of the model object 23 here. In addition, if it is in an example, I and J are set to proportionality constant =1.

[0047] Moreover, the object domain detection means 21 may catch change of the location of the hot object 13 acquired with the infrared camera (infrared image input means) 2, may detect the inspection object domain 20, responds to the magnitude and

the height information of the inspection object domain 20 on a hot object 13 which were further acquired with the infrared camera (infrared image input means) 2, and even if it makes the dimension of the model object 23 expand or reduce, it is not cared about.

[0048] In addition, when it detects that an evaluation result is human being 30, the image field obtained by carrying out a zoom in the 1st field 31 direction of the model object 23 is made into 20 Ns of new inspection object domains. 23 Ns of new model objects of the multiplet structure (what divided the 1st field section into the still finer field (thing equivalent to a face or hair) etc.) are fitted to the 1st field section 31. Face section 34 detection of human being 30 is easily performed by evaluating by performing the same operation as the above-mentioned human detection, and the image near the 23N location of model objects which hits the face section 34 further is recorded on the video tape (recording apparatus) with the date and the sampled data of time of day.

[0049]

[Example] The example of this invention is explained to a detail below, referring to a drawing. Drawing 1 - drawing 11 show the 1st example of this invention. The image processing system 1 using the thermal imagery shown in drawing 1 is equipped with an infrared camera 2, CCD camera 3, and image processing system 1a.

[0050] While the input student image 10 is inputted into this image processing system 1a from the CCD camera (image pick-up means) shown by 3 of drawing, thermal imagery is inputted from the infrared camera (infrared image input means) shown by 2 of drawing. And the object domain detection means 21 has determined the field to inspect from this thermal imagery. The operation using the brightness value of each pixel is performed by the operation means 24 for every field drawn from three persons of the field which the model object 23 established beforehand on the other hand, the input student image 10, and the object domain detection means 21 show, and the evaluation means 25 is performing evaluation processing of whether to be human being according to the result.

[0051] The case where CCD camera 3 samples now the input student image (input image which has brightness gradation) shown by 10 of drawing 2 is considered. Here, 30 is human being and 40 is the lighting under lighting. Since outcrop and lighting 40 of human being 30 emit heat, when they make the image of an infrared camera 2 binary, they will obtain the thermal imagery shown by 12 of drawing 3.

[0052] in addition, the thing used for detection of the object domain which does not perform a direct image processing and is shown the account of the bottom from this

thermal imagery here although an infrared camera 2 tends to be influenced by ambient temperature and the environment and a setup of the threshold level of binary-izing for image analysis is originally difficult -- it is -- threshold level \*\*\*\*\* -- what is necessary is just to set it as low level That is, it is considering as the object of inspection of all hot objects.

[0053] Then, the object domain detection means 21 has set two inspection object domains shown by 20 and 20 to the thermal imagery shown by 12 of drawing 4 from 30 of drawing 3 , and the information on the thermal imagery 12 shown by 40. The field of 20b shown by the dotted line of drawing is a field which considered that the data of 30a were a head and set them up, and the field which the dotted-line field of 20c considered that the data of 30b were a hand, and was set up, and also the dotted-line field of 20a are fields which it considered that 30a was a hand, they considered that 30b was a guide peg, and were set up, and have determined the 1st inspection object domain 20 with the field group of these 20a, 20b, and 20c On the other hand, it considered that the 20d dotted-line field of drawing could not consider 40a which becomes area size other than a head, and has determined as a 2nd inspection object domain 20 with the field similarly shown by 20d.

[0054] In addition, although factors (a distant place, the method of \*\*, an adult, child), such as the model extension of field and contraction, were excluded and detection of this inspection object domain explained on account of explanation if the above-mentioned thermal imagery had it, originally it asks for the area of thermal imagery, and performs a model object also in consideration of expansion or contraction.

[0055] Below, a model object is explained, referring to drawing 6 - drawing 8 . 23 shown in drawing 6 shows a model object, and 23 is divided into two fields which stand in a row in the vertical direction shown by 31 and 32 of drawing. And each field is divided into the field of the double structures 31a and 31b, and the field of 32a and 32b.

[0056] This model object 23 is what modeled human being's configuration, and models head 31 and soma 32 of the human being 30 who shows drawing 7 .

[0057] the dimension which the model object 23 shown in drawing 8 shows the situation at the time of detecting human being 30, and is located in the maximum upper part -- smallness -- the field which the field shown by 31a of drawing among fields 31 shows a man's 30 head, and is shown by 31b shows the circumference background of a head. Moreover, the field which the field shown by 32a of drawing among the fields 32 where a lower part becomes dimension size shows a man's 30 soma, and is shown by 32b shows the circumference background of a soma.

[0058] If it is in the field of 31, each field of the model object 23 is the inside field area of 31a, and the outside of 31a respectively, and it is prepared here so that the inside field area of 31b may carry out abbreviation coincidence. If it is in the field of another side 32, it is the inside field area of 32a, and the outside of 32a respectively, and it is prepared so that the inside field area of 32b may carry out abbreviation coincidence.

[0059] The model object 23 established beforehand as mentioned above is put on the inspection object domain shown by 20 of drawing 4, and 20, and a pixel operation is performed by the operation means 24.

[0060] The operation approach is explained referring to drawing 8 and drawing 9 here. In addition, although an input image originally has a noise in a background and only an object does not lose touch with a background, the image shown in drawing 9 is explained using the easy image pick-up image divided into the part with high brightness (black), and the part with thin brightness (white) on account of explanation. therefore -- although the gradation of each pixel will originally turn into 1024 gradation if a brightness value is a CCD camera with the gradation of 10Bit -- here -- the brightness value of one pixel -- 1K= -- it simplifies to 1024 (black) and 0 (white), and explains below.

[0061] The sum total of the brightness value of each pixel in field 31a shown in drawing 8 is set to Si. Moreover, it is the outside of 31a and the sum total of the brightness value of each pixel in 31b is set to Sj. here -- the area of Si, and the area of Sj -- \*\*\*\* -- equal (100 pixels of each) -- \*\* -- it thinks.

[0062] as the addition result of a brightness value shows temporarily drawing 9, it is referred to as Si=900K and Sj=200K in the soma field (1000 pixels of each) shown by Si=90K, Sj=0, and 32 of drawing in the head field shown by 31 of drawing.

[0063] The operation means 24 is calculating the absolute value of the difference ( $F = |J * \sigma S_j - I * \sigma S_i| * Z$ ) of the brightness value sum total ( $J * \sigma S_j$ ) of the pixel applicable to the outside field Sj of the double structure field which constitutes the model object 23, and the brightness value sum total ( $I * \sigma S_i$ ) of the pixel applicable to the inside field Si of the model object 23 here.

[0064] In addition, if it is in an example, I and J are set to proportionality constant =1 (the same area). Moreover, Z is a weight value which shows the priority of each field. Here, it is calculating by amending Z= 10 in a head field (100 pixels), and amending both surface ratio (pixel number ratio) as Z= 1 in a soma field (1000 pixels).

[0065] This result of an operation becomes  $F(31) * 10 = 900K$  in the head field 31, and becomes  $F(32) * 1 = 700K$  in the soma field 32, and both total F is set to  $F = 900K + 700K = 1600K$ .

[0066] Performing such an operation, it dies in piles to the narrow field (inspection object domain) to which the model object 23 was restricted regularly, and is asking for the location of the model object with which the result of an operation F takes the biggest value.

[0067] The inspection zones 20 and 20 shown in drawing 4 by performing the above operations acquire the biggest F value (result of an operation) in the model object location shown in drawing 5.

[0068] The evaluation means 25 compared reference-value  $R=1000K$  and the result of an operation which are set up beforehand, the thing of the model object location which shows  $F=1600K (>R)$  was judged to be human being, and the image in the direction of the lighting which performed the same operation is judged not to be human being from  $F=75K (<R)$  here.

[0069] Thus, it can be judged that it is the same as human being checks an object by the eye that Background  $S_j$  and the difference of Object  $S_i$  are large, there is a thing equivalent to the head in the field of 31 since two,  $F(31) * 10$  and  $F(32) * 1$ , show the value big [ both ] of 900K and 700K, and there is a thing equivalent to the body in the field of 32.

[0070] In addition, if it is in the above-mentioned example, the multiplet-structure field of the model object 23 is prepared so that the field of the shape of each ellipse may gather at the point shown by K of drawing 6, and the vertical field is carrying out abbreviation connection. moreover, the upper part -- a dimension -- smallness -- it is 2 ream field where two fields (31 32) where it has a field 31 and magnitude differs stand in a row. Furthermore, each of this field has structure of a duplex (the inside field  $S_i$  and outside field  $S_j$ ) respectively. (Ellipse 2 ream duplex structure 22)

[0071] On the other hand, the model object 23 shown in drawing 10 is the model object 23 which is equipped with the 1st field 31 up and stands in a row up and down like the above. Therefore, the configuration of drawing 10 and drawing 11 has the place which is common in it of drawing 6 and drawing 9, gives the same sign to an intersection in drawing 10, and omits detailed explanation.

[0072] The lower part field 32 is established in dimension size, 2 \*\*\*\*s of the double structure field (the 2nd field 32) are made still more equal by abbreviation with the vertical direction, and the model object 23 shown in drawing 10 here is divided into drum section 32A (field of the 2nd A), and leg 32B (field of the 2nd B) in the soma 32 (the 2nd field). That is, it has 3 ream duplex structure 22 with three fields.

[0073] In addition, when the sum total of  $S_i$  shown in drawing 11 and the brightness value which becomes  $S_j$  is searched for, if each weight value =1 of each field (500



pixels) of the 2nd A (23A), and each field (500 pixels) of the 2nd B (32B) As shown in drawing 11 , the result of an operation of each field in the 1st field 31 by  $F(31) = Si(90K) - Sj(0) = 90K * 10$  and 2nd field 32A  $F(32A) = (Si(490K) - Sj(180K)) * 1 = 310K$ , It is set to  $F(32B) = (Si(390K) - Sj(30K)) * 1 = 360K$  by 3rd field 32B, and it is total, and is set to  $F = 1570K$  ( $< R = 1000K$ ), and it is estimated that he is human being.

[0074] Next, the 2nd example of this invention is explained with reference to drawing 12 - drawing 15 .

[0075] The 2nd example arranges CCD camera 3 of the 1st example in the stage in which a neck swing is possible, and it provides it so that the zoom image of a desired image location can be sampled. and a model object is boiled and changed into the thing of the double structure shown in drawing 13 in response to the evaluation result of the 1st example of the above, and the description is in the point of performing the same evaluation as \*\*\*\*. That is, drawing 13 changes the upper field (model object of the maximum upper part) 31 into 23 Ns of new model objects from K necks which show drawing 6 and the model object 23 of drawing 10 to drawing 6 and drawing 10 .

[0076] Therefore, since the configuration of others of the 2nd example is common in it of the 1st example, in drawing 12 - drawing 15 , the same sign is given to an intersection, and detailed explanation is omitted.

[0077] When 20 Ns of inspection object domains shown in drawing 12 detect that the evaluation result shown in the 1st example is human being 30, they are an image field obtained by carrying out a zoom in the 1st field 31 coordinate location of the model object 23.

[0078] And the double structure field shown by 23 Ns of drawing 13 is used as a new model object. 31a of drawing is equivalent to a head, and 31b is a field equivalent to the background around a head.

[0079] A pixel operation is performed while 23 Ns of this model object move in XY side which shows 20 Ns of inspection object domains shown in drawing 14 by the arrow head of drawing.

[0080] 23N location of model objects shown in drawing 12 shows the coordinate location which is 23Ns of model objects with which biggest F was obtained from the result of an operation by the pixel operations Si and Sj. This location where detection of human being was already performed and biggest F was obtained turns into optimal location which catches a head 31.

[0081] \*\*\*\* detection of human being 30 can be easily performed by evaluating by dividing 23 Ns of model objects into a still finer field (thing equivalent to a face or hair) etc., and on the other hand, performing the same operation as the above-mentioned

head detection.

[0082] As for 23 Ns of model objects shown in drawing 15 , the face and 34b of 34a of drawing are hair and a field where 34c shows the background of a head respectively.

[0083] then , if it evaluate by calculate with 23 Ns of model objects by drawing 15 which detect the face section show by 34 of drawing 12 , the detection timing can become what caught people face 34 , and can make an input zoom image record on a video tape ( store ) with the date and the sampled data of time of day .

[0084] In addition, although the model object 23 was made elliptical in this example, you may be the multiplet-structure object of the rectangle which is shown in drawing 17 and which stands in a row up and down. Moreover, the model object 23 may change in the three-step field connected by two places, as shown in drawing 16 and drawing 17 . In addition, the model object 23 shown in drawing 16 and drawing 17 stands in a row in the shape of three step, and divides [ which was prepared like ] into a head 31 (the 1st field), a soma 32 (the 2nd field), and a foot 33 (the 3rd field) the model object field which magnitude differs and stands in a row in the abbreviation vertical direction.

[0085] Although it is what furthermore fixed model object 23 dimension on account of space in this example and being explained, it may respond to magnitude, or the height information or the amount of infrared radiation of the inspection object domain 20 of a hot object 13 obtained with the infrared camera 2, and the dimension of the model object 23 may be made to expand or reduce. Moreover, the object domain detection means 21 is detecting the inspection object domain 20 which catches change of the location of the hot objects 30 and 40 acquired with the infrared camera 2 shown in drawing 3 , and is shown in drawing 4 , and can also remove lighting 40 from a subject of examination beforehand.

[0086] On the other hand, the model object of the \*\*\*\* detection shown in the 2nd example may be the multiplet structure which consists of fields which do not limit only to the configuration shown in the above-mentioned example, and the individual exception, such as opening, and an eye or a lug, left. That is, this invention is not limited to the above-mentioned example, and based on the meaning of this invention, various deformation is possible for it and it does not eliminate these from the range of this invention.

[0087]

[Effect of the Invention] According to this invention, while narrowing down an inspection object domain at high speed using thermal imagery, the image processing system which detects human being correctly can be offered, without being confused by an operating condition, people's appearance, etc.

[0088] Moreover, the image processing system which catches the face zoom image for specifying detected human being is realizable.

---

[Translation done.]

**THIS PAGE BLANK (USPTO)**

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-6642

(P2003-6642A)

(43) 公開日 平成15年1月10日 (2003.1.10)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード(参考)
G 0 6 T 7/00	3 0 0	G 0 6 T 7/00	3 0 0 E 5 B 0 5 7
1/00	3 4 0	1/00	3 4 0 A 5 C 0 5 4
7/20		7/20	3 4 0 B 5 L 0 9 6
H 0 4 N 7/18		H 0 4 N 7/18	A
			D

審査請求 未請求 請求項の数 9 O L (全 10 頁) 最終頁に続く

(21) 出願番号 特願2001-189368(P2001-189368)

(22) 出願日 平成13年6月22日 (2001.6.22)

(71) 出願人 591083532

清水 俊彦

大阪府三島郡島本町江川1丁目15番16-305

(71) 出願人 597163854

見浪 護

福井県福井市文京5-13-7 上里宿舍5-41

(72) 発明者 見浪 護

福井県福井市文京5-13-7 上里宿舍5-41

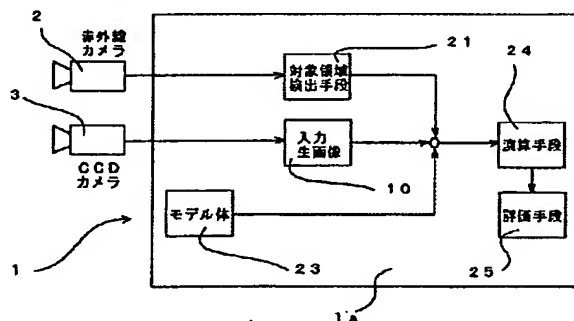
最終頁に続く

(54) 【発明の名称】 熱画像を利用した画像処理装置

(57) 【要約】

【課題】 熱画像を利用して高速で検査対象領域を絞り込むと共に、使用状況や人の容姿などに惑わされることなく正確に人間の検出をおこなう。

【解決手段】 熱画像12をサンプリングする赤外線カメラ2と略同一領域内の可視画像10をサンプリングするCCDカメラ3とを有し、赤外線カメラ2で捕らえた熱画像12から熱体位置の検査対象領域20を求める対象領域検出手段21と、多重構造領域22で構成されて画像対象物を簡易モデル化したモデル体23と、CCDカメラ3による入力生画像10の上記検査対象領域20内に複数のモデル体23が重ねられて、生画像10を用いて検査対象領域20内で画素演算を施す演算手段24と、上記演算結果を評価する評価手段25とを具備する。



## 【特許請求の範囲】

【請求項1】 熱画像を利用して人間検出をおこなう画像処理装置であって、

熱画像をサンプリングする赤外線画像入力手段と略同一領域内の可視画像を撮像する撮像手段とを有し、

赤外線画像入力手段で捕らえた熱画像から熱体位置の検査対象領域を求める対象領域検出手段と、

多重構造領域で構成されて画像対象物を簡易モデル化したモデル体と、

撮像手段で捕らえた入力生画像内の上記検査対象領域へ該モデル体が重ねられて、その領域内の画素輝度値を用いて画素演算を施す演算手段と、

上記演算結果を評価する評価手段とを具備して、人間の検出をおこなうことを特徴とする熱画像を利用した画像処理装置。

【請求項2】 モデル体が、大きさが異なり略上下方向へ連なりかつ2重構造領域で構成されて成る請求項1記載の熱画像を利用した画像処理装置。

【請求項3】 上下に連なるモデル体の下方領域が寸法大に設けられ、その2重構造領域が更に上下方向に2分割されて成る請求項2記載の熱画像を利用した画像処理装置。

【請求項4】 大きさが異なり略上下方向へ連なるモデル体が3段状に連なる請求項2記載の熱画像を利用した画像処理装置。

【請求項5】 演算手段が、モデル体の外側領域に該当する画素の輝度値合計と、モデル体の内側領域に該当する画素の輝度値合計との差の絶対値を求める請求項1、又は2、又は3、又は4記載の熱画像を利用した画像処理装置。

【請求項6】 対象領域検出手段が、赤外線画像入力手段で得られた熱体の位置の変化を捕らえて検査対象領域の検出をおこなう請求項1、又は2、又は3、又は4、又は5記載の熱画像を利用した画像処理装置。

【請求項7】 赤外線画像入力手段で得られた熱体の検査対象領域の大きさに応じて、モデル体の寸法を拡大或いは縮小して成る請求項1、又は2、又は3、又は4、又は5、又は6記載の熱画像を利用した画像処理装置。

【請求項8】 評価結果が人間であることを検出した際、モデル体の最上方領域部方向へズームして得られる画像領域を新たな検査対象領域とし、更に多重構造領域の新たなモデル体で、人間の顔部検出をおこなう請求項1、又は2、又は3、又は4、又は5、又は6、又は7記載の熱画像を利用した画像処理装置。

【請求項9】 評価結果が人間或いはその顔部であることを検出した際、撮像画像を記録する請求項1、又は2、又は3、又は4、又は5、又は6、又は7、又は8記載の熱画像を利用した画像処理装置。

【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、熱画像を利用して人間の検出をおこなう画像処理装置に関するものである。

## 【0002】

【従来の技術】従来から一般的に画像処理装置は、CCDカメラでサンプリングした入力生画像（輝度階調を有する入力画像）を2値化して、細線化処理等を施し端点や分岐点或いは形状などの幾何学的な特徴点を抽出して物体の画像認識をおこなわせるものが知られているが、この場合2値化処理はスレッショールドレベルとの比較で、データは0か1かの2値へ丸められるなど本来画像自身が持っている情報を欠落させてしまうことから幾何学的に規定される形状物でなく、容姿を自在に変えて動く人間などの認識には、特徴点の検出が困難であり不向きであった。

【0003】そこで、進入口などに配されたCCDカメラによる監視画像の変化分を画像処理にて捕らえて、特徴点を抽出して人間らしきものの検出をおこない、更に赤外線検出器を併用して特徴点が人間であると判断させる方法（特開2000-76521号公報、特開2000-155177号公報、特開2000-209573号公報など）が提案されている。

【0004】一方近年、2値化処理を施さずGA（遺伝アルゴリズム）的な探索法により、画像認識をおこなう特開2001-92972号公報などの方法が知られている。

【0005】この特開2001-92972号公報の画像認識方法は、予め判明している対象物の輝度分布や簡易形状をモデルとする探索モデルを入力画像に分布させて、GAを用いて探索モデルの位置を示す遺伝子情報を進化させることで、対象物の位置情報を出力するものである。

【0006】このGAによる遺伝子の進化の過程は、探索モデルの輝度値の分布と入力画像の輝度値の分布とが一致すると相関値が高くなる（一致位置でピークを示す）原理を利用した相関関数を用いて、相関値が高くなる位置へと探索モデルの遺伝子（位置）情報を進化させて解を得ようというものであった。

【0007】なお上記従来の技術においては、2値化画像から動画像の特徴点を抽出するか、GAによって探索モデルを進化させてゆくかの処理をおこなうものであった。

## 【0008】

【発明が解決しようとする課題】上記従来の2値化画像から動画像の特徴点を抽出する構成のものにおいては、照明などの状態でスレッショールドレベルが影響を受けることから外乱に弱く、2値化による情報の丸めで本来もつべき情報を失ってしまう危険がある。また太陽光などが間接的でも差し込む（昼間と夜間で光量に変化す

る)、人の出入りの多いロビーなどで、しかも化粧や髪型服装など容姿の異なる動いている人間の特徴点を抽出するのは至難の業である。

【0009】したがって、特開2000-76521号公報、特開2000-155177号公報、特開2000-209573号公報などのものにあつては、実質上何か物が動いたという動画情報と赤外線検出器の出力をもって人間とみなすものであり、必ずしも正確に人間を検出しているとはいえない。

【0010】しかもこのような動画を抽出する方法にあつては、人が持つカバンなどを振った際、カバンだけの動画部のみが優先する特徴点として抽出されてしまい判断されることから、画像内に人間が存在するにもかかわらずその検出ができないという可能性があり、人間検出をおこなったにも関わらず、検出できなかった(すなわち人はいない)、という重大なミスを犯す問題が指摘されている。そこで、それを回避するために、動く物に加えて赤外線検出器の出力で人間の有無を判断させるものであった。

【0011】したがって利用範囲としては、ある定まった固定背景のエリア(進入口)であつて、人の進入以外考えられない所で、画像の変化(動画)が捉えやすい状況でのみ利用可能な画像処理方法であり、他方例えばCCDカメラ自身が首を振り可動範囲内の画像をサンプリングするような用途のものや、ノイズとなる照明の変動や人の出入りの多い店内やスーパーなどに設置されるような用途、或いは正確に人間であることを検出せねば人命に関わる用途などには対応できないという問題点がある。

【0012】一方、GAによる探索をおこなう上記従来の構成のものにおいては、固有の対象物にあつては、2値化による情報の丸めがなく外乱に強く、しかもパターン一致による比較のものとは異なり、所在の不明な対象物を探索できるメリットがあり、従来のものと比較して有効な方法であつた。

【0013】しかしながら、この探索モデルで探索をおこなう方法のものは、輝度値の分布が判明している探索モデルを予め用意する必要があり、人間を検出するには問題点を有するものであった。

【0014】すなわち、特定の個人の容姿が予め判明しているならば、それは探索可能であるが、一般的な不特定の人間の検出をおこなう場合は、人は肌の色や髪の色が人種により異なり、しかも化粧をして尚かつ、様々な服装をまとうことから、その輝度値の分布をモデル化することは到底困難である。

【0015】しかも、2値化による細線化や特徴点抽出によるもの同様に、探索には多くの時間がかかってしまうという欠点があり、多くの人間を素早く捕らえるには問題がある。

【0016】なお、特開2001-92972号公報に

においては、リアルタイム性を持たせるために、進化の収束を待つことなく所望の制御タイミングで最も高い相関値を持つものを最適解とする例が示されているが、この方法にあつては正確に人間の位置を特定することは言うまでもなく困難である。

【0017】これは、探索モデルのバラまき方や進化の過程に偶然性に依存する要因があり、解の候補とはなつても、必ずしも所望の制御タイミングで最も高い相関値を持つものが真の解とはならないためである。

【0018】更に、画像内には対象物外のものであつても、相関値が中位ピークを示す輝度値分布の場所も数多く存在するため、一時的には誤った領域へGAが注視動作をおこなってしまう可能性もあり、確実にしかも高速処理をおこなわせるには問題となつていた。

【0019】そこで人間検出をおこなう画像処理装置としては、使用状況や人の容姿などの要因に依存することなく、しかも高速で確実に検出するものが望まれていた。

【0020】本発明は、熱画像を利用して高速で検査対象領域を絞り込むと共に、使用状況や人の容姿などに惑わされることなく正確に人間の検出をおこなう画像処理装置を提供することを目的とする。

【0021】

【課題を解決するための手段】本発明の熱画像を利用した画像処理装置は、上記従来例の問題点を解決するため、熱画像を利用して人間検出をおこなう画像処理装置であつて、熱画像をサンプリングする赤外線画像入力手段と略同一領域内の可視画像を撮像する撮像手段とを有し、赤外線画像入力手段で捕らえた熱画像から熱体位置の検査対象領域を求める対象領域検出手段と、多重構造領域で構成されて画像対象物を簡易モデル化したモデル体と、撮像手段で捕らえた入力生画像内の上記検査対象領域へ該モデル体が重ねられて、その領域内の画素輝度値を用いて画素演算を施す演算手段と、上記演算結果を評価する評価手段とを具備して、人間の検出をおこなうことを特徴とする。

【0022】またモデル体が、大きさが異なり略上下方向へ連なりかつ2重構造領域で構成されて成ることが好適である。

【0023】しかも上下に連なるモデル体の下方領域が寸法大に設けられ、その2重構造領域が更に上下方向に2分割されて成るか、又は大きさが異なり略上下方向へ連なるモデル体の領域が3段状に連なるかの何れかにあつては、人の容姿に依存することなく的確な検出をおこなうものとなる。

【0024】そして演算手段が、モデル体の外側領域に該当する画素の輝度値合計と、モデル体の内側領域に該当する画素の輝度値合計との差の絶対値を求めることが、簡単な演算で背景との差を生じるものとなる。

【0025】また、対象領域検出手段が、赤外線画像入

力手段で得られた熱体の位置の変化を捕らえて検査対象領域の検出をおこなうものが好適である。

【0026】さらに赤外線画像入力手段で得られた熱体の検査対象領域の大きさに応じて、モデル体の寸法を拡大或いは縮小して成ることで、子供や大人、或いは遠方や近傍の検出をおこなうことができる。

【0027】なお、評価結果が人間であることを検出した際、モデル体の最上方領域部方向へズームして得られる画像領域を新たな検査対象領域とし、更に多重構造領域の新たなモデル体を適応すれば、人間の顔部検出をおこなうことができる。そして、その撮像画像を記録するものであれば、その人物を特定する画像を得るものとなる。

【0028】（作用）本発明は上記構成によって、次のような作用を営むことができる。すなわち、赤外線画像入力手段で捕らえた熱画像から熱体位置の検査対象領域を求めることで、熱画像は人間（正確には顔や手など肌が露出している部分）、動物、発熱体、蓄熱体などに反応し、人間の存在らしき位置の検出を画像探索をおこなうことなく、確実にしかも超高速で把握して捕らえることができ、人間の大きさを考慮して発熱体の分布状況から幾つかの検査対象領域を素早く設定することができる。

【0029】一方、多重構造領域で構成されて画像対象物を簡易モデル化したモデル体は、対象物の形状を簡易的にモデル化し、更に多重構造（例えば2重構造）となる領域を備えている。そして、この多重構造モデルと重なる画素の輝度値から人間であることを判断させる為に、2重構造の内側の領域を人間そのものに対応させるものであって、外側の領域が背景に相当するものを示すものとなっている。なお、上記発熱体の検出により従来例で示すようなGAなどの探索モデルを用いる探索の必要がないことから、このモデル体には対象物の輝度値の分布データは不要であり、任意輝度値の均一データであってもいいし、形状のみを定義する空枠であっても良い。

【0030】そこで、撮像手段による入力生画像に対して、上記求められた検査対象領域内に順次モデル体の空枠が重ねられて、輝度値を有する画素から検査対象領域内で画素演算が施される。このモデル体の重ね作業は、限られた狭い検査領域内にモデル体が挿入可能な配置で順次重ねられて、上述内側（人間）の領域と外側（背景）の領域とで、それぞれの領域内の画素の輝度値が求められ、外側領域（背景）の輝度値と内側領域（人間）の輝度値に差が生じれば、それは人が物を識別するのと同様に、この差をもってモデル体形状に合致する物体の存在を知ることができる。

【0031】この演算結果は、内側の領域が人間の形状を成し、しかもそれが熱体であることから、演算結果の差が際立つものであれば、それは人間であると評価する

ことができる。

【0032】またモデル体が、大きさが異なり略上下方向へ連なりかつ2重構造領域で構成されて成るものであれば、モデル体の上方の領域が示すものを頭部とし、下方の領域が示すものを体部とすることができる。この2つの領域で各々内側領域（頭部）と外側領域（背景）かつ内側領域（体部）と外側領域（背景）の関係で、各領域で輝度値に差を生じるものならば、そこには熱を発するものがあり、しかも頭と体に相当する形状物がある、すなわち人間であると判断することができる。

【0033】しかも上下に連なるモデル体の下方領域が寸法大に設けられ、その2重構造領域が更に上下方向に2分割されるものであれば、分割された上部領域が示すものを胴部とし、下部領域が示すものを脚部とすることができる。この上下2つの領域で上記同様に、内側領域（胴部）と外側領域（背景）かつ内側領域（脚部）と外側領域（背景）の関係で、輝度値に差を生じるものならば、そこには胴と脚があると判断することができ、このようなモデル体であれば人の容姿に依存することなく的確な検出をおこなうモデル体となる。

【0034】一方、大きさが異なり略上下方向へ連なるモデル体の領域が3段状に連なるものにあては、3段に分割された上部領域が示すものを頭部とし、中部領域が示すものを胴部とし、下部領域が示すものを脚部とすることができる。この中下2つの領域で上記2分割領域のもの同様に、内側領域（胴部）と外側領域（背景）かつ内側領域（脚部）と外側領域（背景）の関係で、輝度値に差を生じるものならば、そこには胴と脚があると判断することができ、上記同様に、このようなモデル体であれば人の容姿に依存することなく的確な検出をおこなうモデル体となる。

【0035】なお、上述のように演算手段が、モデル体の外側領域に該当する各画素の輝度値の合計と、モデル体の内側領域に該当する各画素の輝度値の合計との差の絶対値を求めるものであれば、この2重構造が故に簡単な加減算の演算のみで、モデル体が対象物形状と重なる位置で、際立つ（背景との）差を生じさせることができる。

【0036】また、対象領域検出手段が、赤外線画像入力手段で得られた熱体の位置の変化を捕らえて検査対象領域の検出をおこなうものであれば、例えばマネキンなどが熱せられて置かれていたとしても、動かないため検出対象とはならず、同様に動かない発熱体を無視することで、無駄なく、より確実に人間のみを検出することができる。

【0037】さらに赤外線画像入力手段で得られた熱体の検査対象領域の大きさに応じて、モデル体の寸法を拡大或いは縮小して成るものであれば、子供や大人、或いは遠方や近傍のモデルを用いて、より詳しく人間検出をおこなわせることができる。

10

20

30

40

50



【0038】なお、評価結果が人間であることを検出した際、モデル体の最上方領域部方向（頭部）へズームして得られる画像領域を新たな検査対象領域とし、更に多重構造領域の新たなモデル体を適応すれば、限られた最上方領域部周辺のみへ人の頭部や顔部などに相当する新たな多重モデルが適応されて、既に人間であることが検出されているものへ、なおかつズームされた限られた画像領域で、（例えば顔と髪の差などの多重領域をもつモデルで）顔部だけのモデル体を使って、上記同様に輝度値による画素演算を施し評価をおこなうことができる。故に入力画像全てを探索サーチさせることなく、限られたズーム領域のみで対象となる顔が撮像手段の方向に向いた際、その顔部のズーム画像を捕らえることが可能となる。

【0039】更に上記顔部を捕らえるタイミングで、顔部モデル体位置近傍の撮像画像を逐次記録するものであれば、その人物を特定する画像を得るものとなり、人間検出のみならず人物特定に寄与する画像の記録をおこなうことができる。

【0040】このような本発明の熱画像を利用した画像処理装置は、使用状況や人の容姿などの要因に依存することなく、しかも高速で確実に人間を検出することができる。

【0041】また、検出された人間を特定するための顔面ズーム画像を捕らえる画像処理装置を実現することができる。

【0042】

【発明の実施の形態】本発明の実施の形態について図1～図17を参照して簡単に説明する。

【0043】本発明の熱画像を利用した画像処理装置1は、赤外線カメラ2で捕らえた画像（熱画像）を利用して人間30の検出をおこなう画像処理装置1aであって、熱画像12をサンプリングする赤外線カメラ（赤外線画像入力手段）2と略同一領域内の一部又は同一の可視画像（可視光の画像）10を撮像することのできる光学式白黒のCCDカメラ（撮像手段）3とを有し、赤外線カメラ（赤外線画像入力手段）2で捕らえた熱画像12を2値化して一定温以上の熱体物13の座標位置から検査対象領域20の大きさを決める対象領域検出手段21と、2重構造状の楕円領域22で構成されて画像対象物の形状を予め簡易的にモデル化したモデル体23と、CCDカメラ（撮像手段）3による入力生画像（輝度階調を有する入力画像）10の該検査対象領域20内に上記モデル体23が規則的に（或いは不規則に）重ねられて、モデル体23の示す領域内の画素の輝度値から画素演算を施す演算手段24と、上記演算結果が予め設定された値に達するか否かで評価する評価手段25とを具備して、人間30の検出をおこなっている。

【0044】またモデル体23が、人30の首部に当たる点（K点）で交わり、頭部31に相当する寸法小なる

第1の領域と体部32に相当する寸法大なる第2の領域で構成されることで、大きさが異なり略上下方向へほぼ連なり、かつ上記一点（K点）に寄り合うように2重の楕円構造領域22の構造で構成されている。

【0045】しかも上下に連なるモデル体の下方領域が寸法大に設けられ、その2重構造領域（第2の領域32）が更に上下方向に略均等に2分割され、体部32（第2の領域）を胸部32A（第2Aの領域）と脚部32B（第2Bの領域）とに分割している。なお、大きさが異なり略上下方向へ連なるモデル体領域を3段状に連なるものとしても同様に、体部32（第2の領域）と足部33（第3の領域）とに分割できるものとなる。

【0046】ここで演算手段24は、モデル体23を構成する2重構造領域22の外側領域 $S_j$ に該当する画素の輝度値合計（ $J * \sum S_j$ ）と、モデル体23の内側領域 $S_i$ に該当する画素の輝度値合計（ $I * \sum S_i$ ）との差（ $|J * \sum S_j - I * \sum S_i|$ ）の絶対値に重み値を乗じて求めている。なお実施例にあってはI、Jは比例定数=1としている。

【0047】また、対象領域検出手段21は、赤外線カメラ（赤外線画像入力手段）2で得られた熱体13の位置の変化を捕らえて検査対象領域20の検出をおこなってもいいし、さらに赤外線カメラ（赤外線画像入力手段）2で得られた熱体13の検査対象領域20の大きさや高さ情報に応じて、モデル体23の寸法を拡大或いは縮小させてもかまわない。

【0048】なお、評価結果が人間30であることを検出した際、モデル体23の第1領域31方向へズームして得られる画像領域を新たな検査対象領域20Nとし、第1領域部31へ多重構造（第1領域部を更に細かな領域（顔や髪に当たるもの）などに分けたもの）の新たなモデル体23Nを適応させて、上記人間検出同様の演算を施し評価をおこなうことで容易に人間30の顔部34検出をおこない、更に顔部34に当たるモデル体23N位置近傍の画像をビデオテープ（記録装置）に日付けやサンプリングした時刻のデータと共に記録している。

【0049】

【実施例】以下本発明の実施例について、図面を参照しながら詳細に説明する。図1～図11は本発明の第1実施例を示すものである。図1に示す熱画像を利用した画像処理装置1は、赤外線カメラ2、CCDカメラ3、画像処理装置1aを備えている。

【0050】この画像処理装置1aには図の3で示すCCDカメラ（撮像手段）から入力生画像10が入力されると共に、図の2で示す赤外線カメラ（赤外線画像入力手段）から熱画像が入力されている。そして、この熱画像から対象領域検出手段21は検査する領域を決定している。一方予め設けられているモデル体23と入力生画像10と対象領域検出手段21の示す領域の3者から導き出される領域毎に、各画素の輝度値を用いた演算が演

算手段24でおこなわれ、その結果に応じて評価手段25は人間であるか否かの評価処理をおこなっている。

【0051】今CCDカメラ3が図2の10で示す入力生画像（輝度階調を有する入力画像）をサンプリングした場合について考える。ここで、30は人間であり、40は点灯中の照明である。人間30の露出部と照明40は熱を発することから、赤外線カメラ2の画像を2値化すると図3の12で示す熱画像を得ることになる。

【0052】なお本来、赤外線カメラ2は周辺温度や環境に左右され易く、画像解析のための2値化のスレッシュホールドレベルの設定は難しいが、ここではこの熱画像から直接画像処理をおこなうものではなく、下記示す対象領域の検出に利用するものであって、スレッシュホールドレベル極めて低いレベルに設定しておけばよい。すなわち、熱体全てを検査の対象としている。

【0053】そこで、図3の30、40で示す熱画像12の情報から対象領域検出手段21は、図4の12で示す熱画像へ20、20で示す2つの検査対象領域を設定している。図の点線で示す20bの領域は30aのデータを頭部と見なし設定した領域であり、20cの点線領域は30bのデータを手と見なし設定した領域、更に20aの点線領域は30aを手、30bを足と見なし設定した領域であり、これら20a、20b、20cの領域群をもって1つ目の検査対象領域20を決定している。一方図の20dの点線領域は面積大なる40aを頭部以外には考えられないと見なし、同じく20dで示す領域をもって2つ目の検査対象領域20として決定している。

【0054】なお、この検査対象領域の検出は、上記熱画像にあっては説明の都合上、モデル体の拡大や縮小などの要因（遠方、近方、大人、子供）を省き説明したが、本来熱画像の面積を求めて、モデル体を拡大或いは縮小をも考慮しておこなうものである。

【0055】次ぎに、図6～図8を参照しながらモデル体について説明する。図6に示す23はモデル体を示すもので、23は図の31と32で示す上下方向に連なる2つの領域に分けられている。そして各領域は2重構造31a、31bの領域と、32a、32bの領域とに分割されている。

【0056】このモデル体23は人間の形状をモデル化したもので、図7に示す人間30の頭部31と体部32とをモデル化している。

【0057】図8に示すモデル体23は人間30を検出した際の様子を示すもので、最上方に位置する寸法小なる領域31のうち、図の31aで示す領域が人30の頭部を示し、31bで示す領域が頭部の周辺背景を示している。また下方の寸法大なる領域32のうち、図の32aで示す領域が人30の体部を示し、32bで示す領域が体部の周辺背景を示している。

【0058】ここでモデル体23の各領域は例えば31

の領域にあっては、各々31aの内側領域面積と31aの外側であって31bの内側領域面積とが略一致するよう設けられている。他方32の領域にあっては、各々32aの内側領域面積と32aの外側であって32bの内側領域面積とが略一致するよう設けられている。

【0059】以上のように予め設けられたモデル体23は、図4の20、20で示す検査対象領域に重ねられて、演算手段24により画素演算がおこなわれる。

【0060】ここで図8、図9を参照しながらその演算方法を説明する。なお、本来入力画像は背景にノイズがあり対象物のみが背景から浮き上がることはないが、図9に示す画像は説明の都合上、輝度の高い部分（黒色）と輝度の薄い部分（白色）とに分けた簡単な撮像画像を用いて説明する。したがって本来輝度値は、10Bitの階調をもつCCDカメラであれば各画素の階調は1024階調となるが、ここでは1つの画素の輝度値を1K=1024（黒）と0（白）とに簡略化して以下説明する。

【0061】図8に示す領域31a内の各画素の輝度値の合計を $S_i$ とする。また31aの外側であって31b内の各画素の輝度値の合計を $S_j$ とする。ここで $S_i$ の面積と $S_j$ の面積はほぼ等しい（各々100画素）と考える。

【0062】仮に輝度値の加算結果が図9に示すように、図の31で示す頭部領域で $S_i=900K$ 、 $S_j=0$ 、また図の32で示す体部領域（各々1000画素）で $S_i=900K$ 、 $S_j=200K$ とする。

【0063】ここで演算手段24は、モデル体23を構成する2重構造領域の外側領域 $S_j$ に該当する画素の輝度値合計（ $J \times \sum S_j$ ）と、モデル体23の内側領域 $S_i$ に該当する画素の輝度値合計（ $I \times \sum S_i$ ）との差（ $F = |J \times \sum S_j - I \times \sum S_i| \times Z$ ）の絶対値を求めている。

【0064】なお実施例にあってはI、Jは比例定数=1（同一面積）としている。またZは各領域のプライオリティを示す重み値である。ここでは、頭部領域（画素数100）で $Z=10$ 、体部領域（画素数1000）で $Z=1$ として、両者の面積比（画素数比）を補正して演算をおこなっている。

【0065】この演算結果は、頭部領域31で $F(31) \times 10 = 900K$ 、体部領域32で $F(32) \times 1 = 700K$ となり、両者のトータルFは、 $F = 900K + 700K = 1600K$ となる。

【0066】このような演算を施しながら、モデル体23を規則的に限られた狭い領域（検査対象領域）へ重ねてゆき、演算結果Fが最も大きな値を取るモデル体の位置を求めている。

【0067】以上のような演算を施すことで、図4に示す検査領域20、20は、図5に示すモデル体位置で最も大きなF値（演算結果）を得る。

10

20

30

40

50

【0068】ここで評価手段25は、予め設定されている基準値 $R=1000K$ と演算結果とを比較して、 $F=1600K(>R)$ を示すモデル体位置のものを人間であると判断し、同様の演算を施した照明の方の画像を $F=75K(<R)$ から人間でないと判断している。

【0069】このように、背景 $S_j$ と対象 $S_i$ の差が大きいということは、人間が目で物を確認するのと同様であり、 $F(31)*10$ と $F(32)*1$ の2つが $900K$ と $700K$ という共に大きな値を示すことから31の領域には頭に相当するものがあり、32の領域には体に相当するものと判断できる。

【0070】なお、上記実施例にあっては、モデル体23の多重構造領域は図6のKで示す点で各楕円状の領域が寄り合うように設けられ、上下領域が略連結している。また上方に寸法小なる領域31をもち、大きさの異なる2つの領域(31、32)が連なる2連領域である。更に、この各領域が各々2重(内側領域 $S_i$ と外側領域 $S_j$ )の構造となっている。(楕円2連2重構造22)

【0071】一方、図10に示すモデル体23は、上方に第1の領域31を備えて、上記同様に上下に連なるモデル体23である。したがって、図10及び図11の構成は図6及び図9のそれと共通するところがあり、図10において共通部分に同一符号を付し詳細な説明を省略する。

【0072】ここで図10に示すモデル体23は、下方領域32が寸法大に設けられ、その2重構造領域(第2の領域32)が更に上下方向に略均等に2分割され、体部32(第2の領域)を胴部32A(第2Aの領域)と脚部32B(第2Bの領域)とに分割されている。すなわち3つの領域をもつ3連2重構造22となっている。

【0073】なお、図11に示す $S_i$ 、 $S_j$ なる輝度値の合計が求められた場合、第2A(23A)の各領域(500画素)及び第2B(32B)の各領域(500画素)のそれぞれの重み値=1とすると、各領域の演算結果は、図11に示すように、第1の領域31で $F(31)=S_i(90K)-S_j(0)=90K*10$ 、第2の領域32Aで $F(32A)=(S_i(490K)-S_j(180K))*1=310K$ 、第3の領域32Bで $F(32B)=(S_i(390K)-S_j(30K))*1=360K$ となり、トータルで $F=1570K(<R=1000K)$ となり人間であると評価されるものである。

【0074】次に、本発明の第2の実施例を図12～図15を参照して説明する。

【0075】第2実施例は第1実施例のCCDカメラ3を首振り可能なステージに配設して、所望の画像位置のズーム画像をサンプリングできるように設けてある。そして、上記第1実施例の評価結果を受けて、モデル体を図13に示す2重構造のものに変えて、上述同様の評

価をおこなう点に特徴がある。すなわち図13は、図6、図10のモデル体23を、図6、図10に示す首部K点より上方の(最上方のモデル体)領域31を新たなモデル体23Nに変更したものである。

【0076】したがって、第2実施例のその他の構成は第1実施例のそれと共通しているもので、図12～図15において共通部分に同一符号を付し詳細な説明を省略する。

【0077】図12に示す検査対象領域20Nは、第1実施例で示す評価結果が人間30であることを検出した際、モデル体23の第1領域31座標位置へズームして得られる画像領域である。

【0078】そして、図13の23Nで示す2重構造領域を新たなモデル体としている。図の31aは頭部に相当し、31bは頭部周辺の背景に相当する領域である。

【0079】このモデル体23Nは、図14に示す検査対象領域20Nを図の矢印で示すXY面を移動しながら画素演算がおこなわれるものである。

【0080】図12に示すモデル体23N位置は、画素演算 $S_i$ 、 $S_j$ による演算結果から最も大きなFが得られたモデル体23Nの座標位置を示すものである。既に人間の検出がおこなわれ、最も大きなFが得られたこの位置は、頭部31を捕らえる最適な位置となる。

【0081】一方、モデル体23Nを更に細かな領域(顔や髪に当たるもの)などに分けて、上記頭部検出同様の演算を施し評価をおこなうことで容易に人間30の顔部検出をおこなうことができる。

【0082】図15に示すモデル体23Nは、図の34aは顔面、34bは髪、34cは頭部の背景を各々示す領域である。

【0083】そこで、図12の34で示す顔部を検出する図15によるモデル体23Nで演算して評価をおこなうものであれば、その検出タイミングは人の顔面34を捕らえたものとなり、入力ズーム画像をビデオテープ(記憶装置)に日付けやサンプリングした時刻のデータと共に記録させることができる。

【0084】なお、本実施例においてはモデル体23を楕円形状としたが、図17に示す上下に連なる長方形の多重構造体であってもかまわない。また、モデル体23は、図16、図17に示すように、2カ所で連結する3段領域で成るものであってもよい。なお、図16、図17に示すモデル体23は、大きさが異なり略上下方向へ連なるモデル体領域を3段状に連なる様に設けたもので、頭部31(第1の領域)と体部32(第2の領域)と足部33(第3の領域)とに分割したものである。

【0085】さらに本実施例においては紙面の都合上モデル体23寸法を固定したもので説明したが、赤外線カメラ2で得られた熱体13の検査対象領域20の大きさや高さ情報或いは赤外線量に応じて、モデル体23の寸法を拡大或いは縮小させてもかまわない。また、対象領

域検出手段 21 は、図 3 に示す赤外線カメラ 2 で得られた熱体 30、40 の位置の変化を捕らえて、図 4 に示す検査対象領域 20 の検出をおこなうことで、予め照明 40 を検査対象から除くことも可能である。

【0086】一方、第 2 実施例に示す顔部検出のモデル体は、上記実施例に示す形状のみに限定するものではなく、口や目或いは耳など個別の離れた領域で構成される多重構造であってもかまわない。すなわち本発明は上記実施例に限定されるものではなく、本発明の趣旨に基づき種々の変形が可能であり、これらを本発明の範囲から

【0087】

【発明の効果】本発明によれば、熱画像を利用して高速で検査対象領域を絞り込むと共に、使用状況や人の容姿などに惑わされることなく正確に人間の検出をおこなう画像処理装置を提供することができる。

【0088】また、検出された人間を特定するための顔面ズーム画像を捕らえる画像処理装置を実現することができる。

【図面の簡単な説明】

【図 1】本発明の実施例を示すブロック図。

【図 2】入力生画像を示す説明図。

【図 3】その熱画像を示す説明図。

【図 4】その検査対象領域を示す概念図。

【図 5】その検査方法を示す概念図。

【図 6】モデル体を示す原理図。

【図 7】検査対象を示す斜視図。

【図 8】その検査方法を示す概念図。

【図 9】その演算方法を示す概念図。

【図 10】その他のモデル体を示す原理図。

【図 11】その演算方法を示す概念図。

【図 12】本発明の第 2 実施例を示す概念図。

【図 13】頭部領域を示すモデル体の原理図。

【図 14】その演算方法を示す概念図。

【図 15】顔部領域を示すモデル体の原理図。

【図 16】その他のモデル体の構成例を示す概念図。

【図 17】その他のモデル体の形状例を示す概念図。

【符号の説明】

1 熱画像を利用した画像処理装置

1 a 画像処理装置

2 赤外線カメラ（赤外線画像入力手段）

3 CCD カメラ（撮像手段）

10 入力生画像（輝度階調を有する入力画像）

12 熱画像

20 検査対象領域

21 対象領域検出手段

22 多重構造領域（2 重構造領域）

23 モデル体

24 演算手段

25 評価手段

30 人間、人

31 頭部（第 1 の領域）

32 体部（第 2 の領域）

32 A 胴部（第 2 A の領域）

32 B 脚部（第 2 B の領域）

33 足部（第 3 の領域）

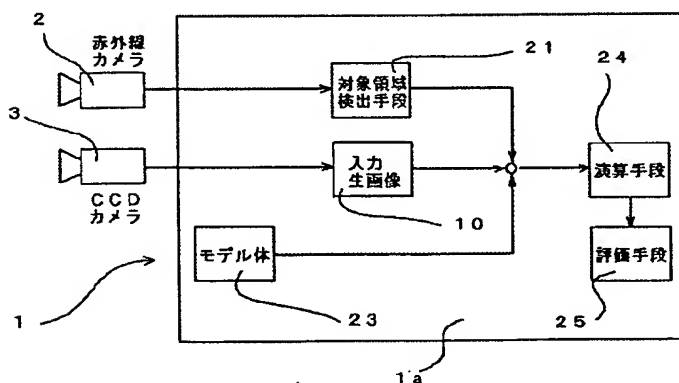
34 顔部

K 首部（K 点）

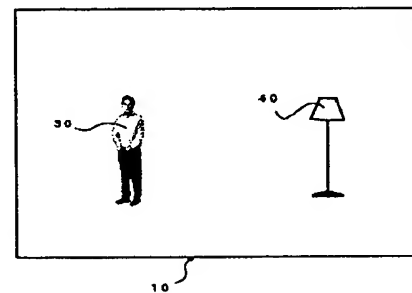
S i 内側領域

S j 外側領域

【図 1】

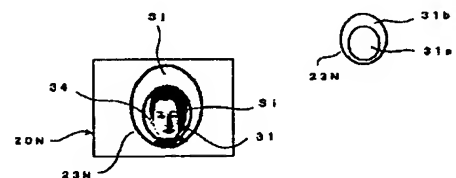


【図 2】

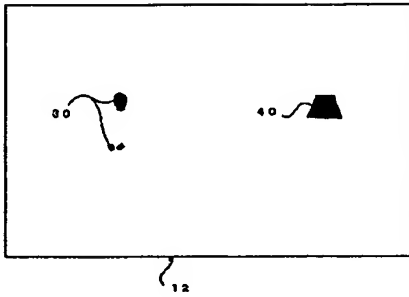


【図 12】

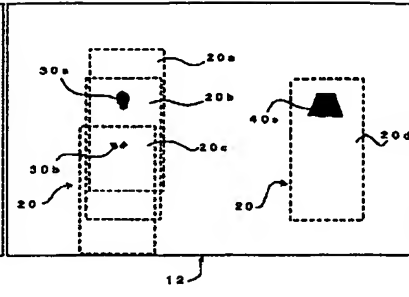
【図 13】



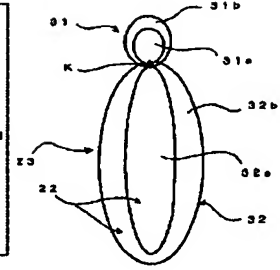
【図3】



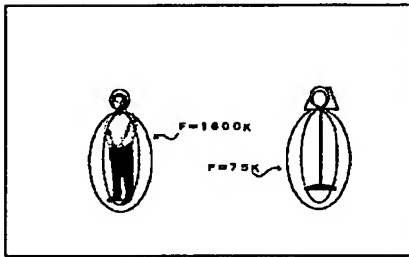
【図4】



【図6】



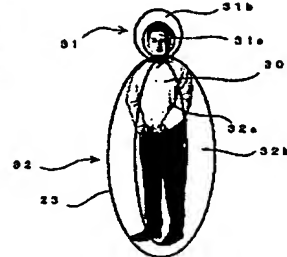
【図5】



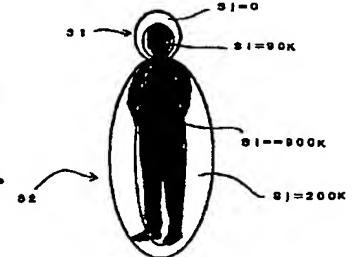
【図7】



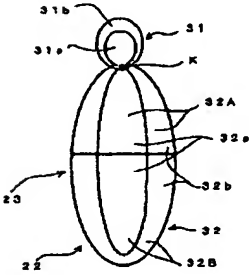
【図8】



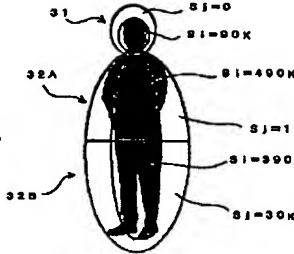
【図9】



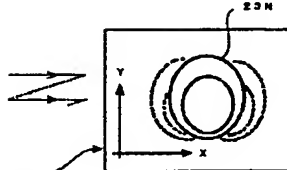
【図10】



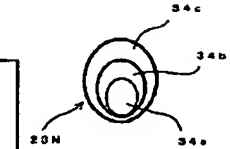
【図11】



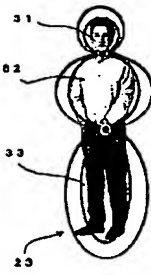
【図14】



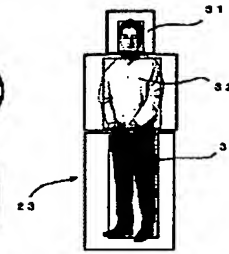
【図15】



【図16】



【図17】



フロントページの続き

(51)Int.Cl.<sup>7</sup>  
H04N 7/18

識別記号

F I  
H04N 7/18

キーワード(参考)

K  
N

(72)発明者 清水 俊彦  
大阪府三島郡島本町江川1丁目15番16-  
305

Fターム(参考) 5B057 AA19 BA02 BA08 BA13 CA08  
CA12 CA16 CD05 CE08 CE09  
DA06 DB02 DB09 DC22  
5C054 AA01 CA04 CA05 FC03 FC11  
FF03 HA18  
5L096 AA06 BA03 CA05 EA03 EA35  
FA69 FA79 HA02